

# USING GENERATIVE AI WITH HUMAN-IN-THE-LOOP TO DELIVER TIMELY, PERSONALIZED FEEDBACK AT SCALE

U.T.F. LAM<sup>1</sup>, R.M. SINGH<sup>1</sup>, K.H.K. YUEN<sup>2</sup>, J.Y. CHEN<sup>2</sup> and F.M. YEONG<sup>1</sup>

<sup>1</sup>Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore

<sup>2</sup>AI Centre for Educational Technologies, School of Computing, National University of Singapore

## BACKGROUND

LSM2233 Cell Biology is a large-enrolment undergraduate course (~200 students) that focuses on developing a comprehensive understanding of subcellular structures, functions, and interactions in both unicellular and multicellular systems. Our objective is to support student learning in constructing clear, evidence-based explanations grounded in experimental data. To this end, we developed practice Short Answer Questions (SAQs) with personalized feedback released in a timely manner to students. Here, we provide an initial analysis of the platform with respect to time savings, the accuracy of AI-generated feedback, and student perceptions of AI-assisted, human-in-the-loop feedback generated by Coursemology platform. We suggest that this approach can increase students' opportunities for meaningful practice and timely feedback while maintaining feedback quality, without substantially increasing instructor workload.

## METHODOLOGY

- 6 SAQs (optional and carried no course credit)
- Questions aligned with the intended learning outcomes for topics of lysosome, cell cycle, secretion and signalling
- Implemented on Coursemology, an LMS previously developed by NUS AICET (Chua et al., 2026)
- Human-in-the-loop workflow (Fig. 1) (Memarian & Doleck, 2024)

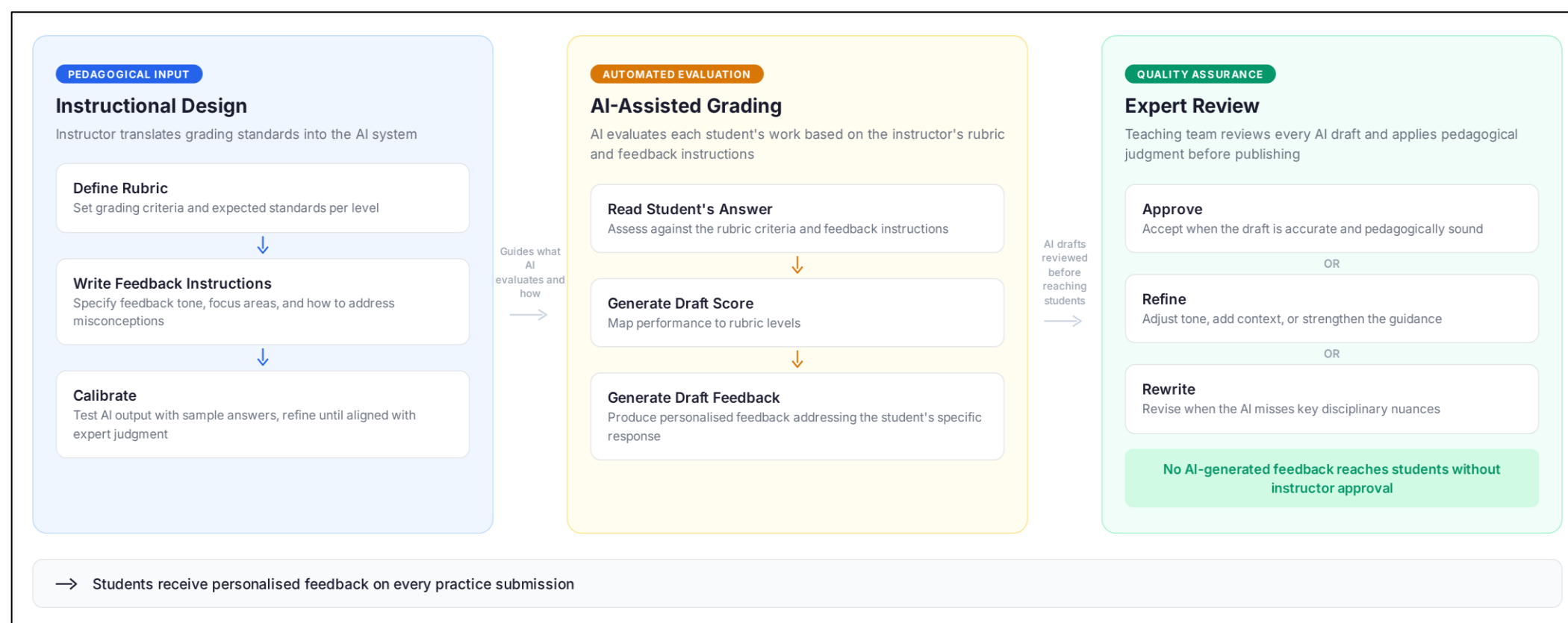


Fig 1. Human-in-the-loop AI-assisted feedback workflow in Coursemology.

Instructors define rubrics, write feedback instructions, and calibrate with sample answers to expert judgement (left). The GenAI reads each submission, generates a draft score and personalized feedback aligned to the rubric (middle). Instructors approve, refine, or rewrite drafts when disciplinary nuances are missed (right). Only instructor-approved feedback is released to students for every attempt to ensure pedagogical quality and accountability.

- Instructor-reviewed AI drafts
- Customised prompts specifying required content, feedback tone, and structure (Fig. 2), together with model answers.
- Iteratively tested and refined against sample responses by instructors, with calibration based on instructors' editing patterns (Fig. 3)

Custom Prompt  
Add grading instructions (e.g. question context, model answer, feedback tone). Leave blank if unsure.

Paragraph AI B I U S X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16 X17 X18 X19 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X30 X31 X32 X33 X34 X35 X36 X37 X38 X39 X40 X41 X42 X43 X44 X45 X46 X47 X48 X49 X50 X51 X52 X53 X54 X55 X56 X57 X58 X59 X60 X61 X62 X63 X64 X65 X66 X67 X68 X69 X70 X71 X72 X73 X74 X75 X76 X77 X78 X79 X80 X81 X82 X83 X84 X85 X86 X87 X88 X89 X90 X91 X92 X93 X94 X95 X96 X97 X98 X99 X100

You are an associate professor of cell biology teaching a cell biology course for second-year undergraduates. Provide feedback to the student's submission by referring closely to the Model answer provided below. Also use the following instructions and carefully provide feedback to the student's submission. The tone of the feedback should be constructive but not overly enthusiastic, especially if the answer is not good or contains any errors. In the feedback, instead of saying "the student's answer", you should address the feedback directly to the student and say "Your response" or "your answer". In the feedback, do not use the term "model" or "model answer" or refer to it as we do not want students to think that there is only one particular answer and we are not providing it to students.

To provide feedback to the student, evaluate each sentence in the student's submission carefully for correctness based on the model answer. If the student makes both a correct and an incorrect assertions in the sentence, classify it as **Wrong** and explain both what was right and what was wrong in the sentence.

For the content in addition to the model answer to be used as reference, note also the instructions in

Fig 2. Example of Coursemology customized prompts tailored made for practice SAQs in LSM2233

- Student responses were exported from Coursemology in JavaScript Object Notation (JSON) format and converted into spreadsheets for analysis
- Mid- and end-semester anonymous surveys (Google Forms) captured students' perceptions of the feedback.

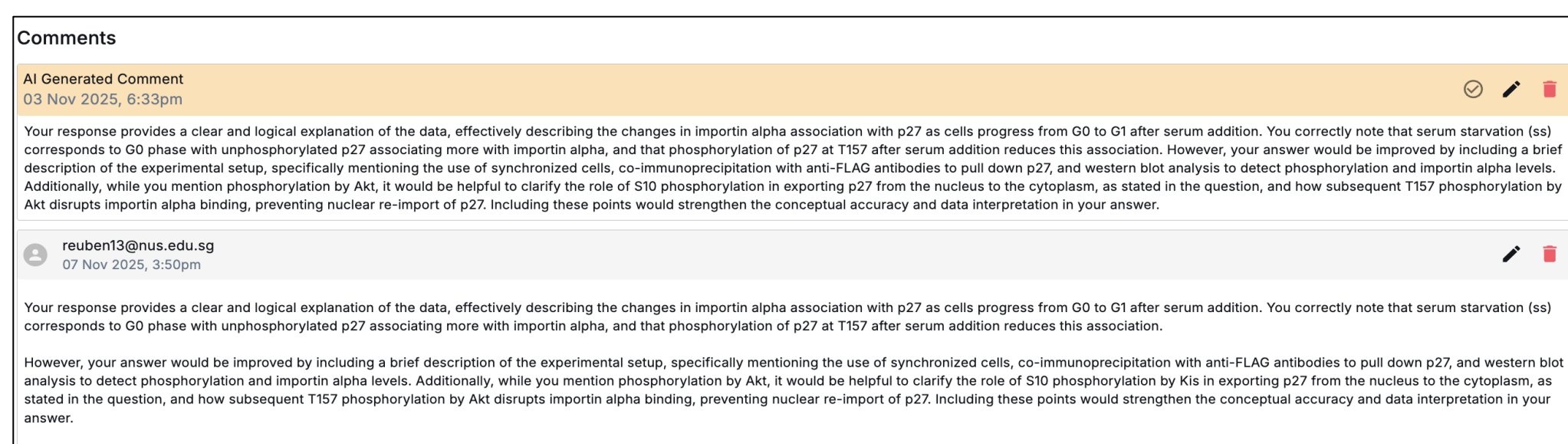


Fig 3. Sample of AI-generated comments (above) from practice SAQ of cell cycle topic in LSM2233 and the instructors' comments for release (below)

## RESULTS

- 561 personalised feedback responses
- Increased opportunities for student practice without increasing instructor workload, critical for large-enrolment classes
- ~5x reduction in feedback time (Fig. 4)
- 88% of AI drafts usable with minor to moderate edits (Fig. 5)
- AI served as a reliable first-pass feedback generator, refined by instructor expertise
- Improved draft quality over time with model upgrades (GPT-4o mini → GPT-4.1 mini)
- Fewer major instructor interventions required as system performance improved

Time spent for	Human instructors only	With Coursemology
Marking	15 min	0 min
Writing feedback	10 min	0 min
Reviewing feedback	0 min	5 min
<b>Total:</b>	<b>25 min</b>	<b>5 min</b>

Fig. 4. Time required for feedback with and without AI support.

Comparison of instructor time per student shows a reduction from ~25 minutes (manual marking and feedback) to ~5 minutes with Coursemology, where time is primarily spent reviewing and editing AI-generated feedback.

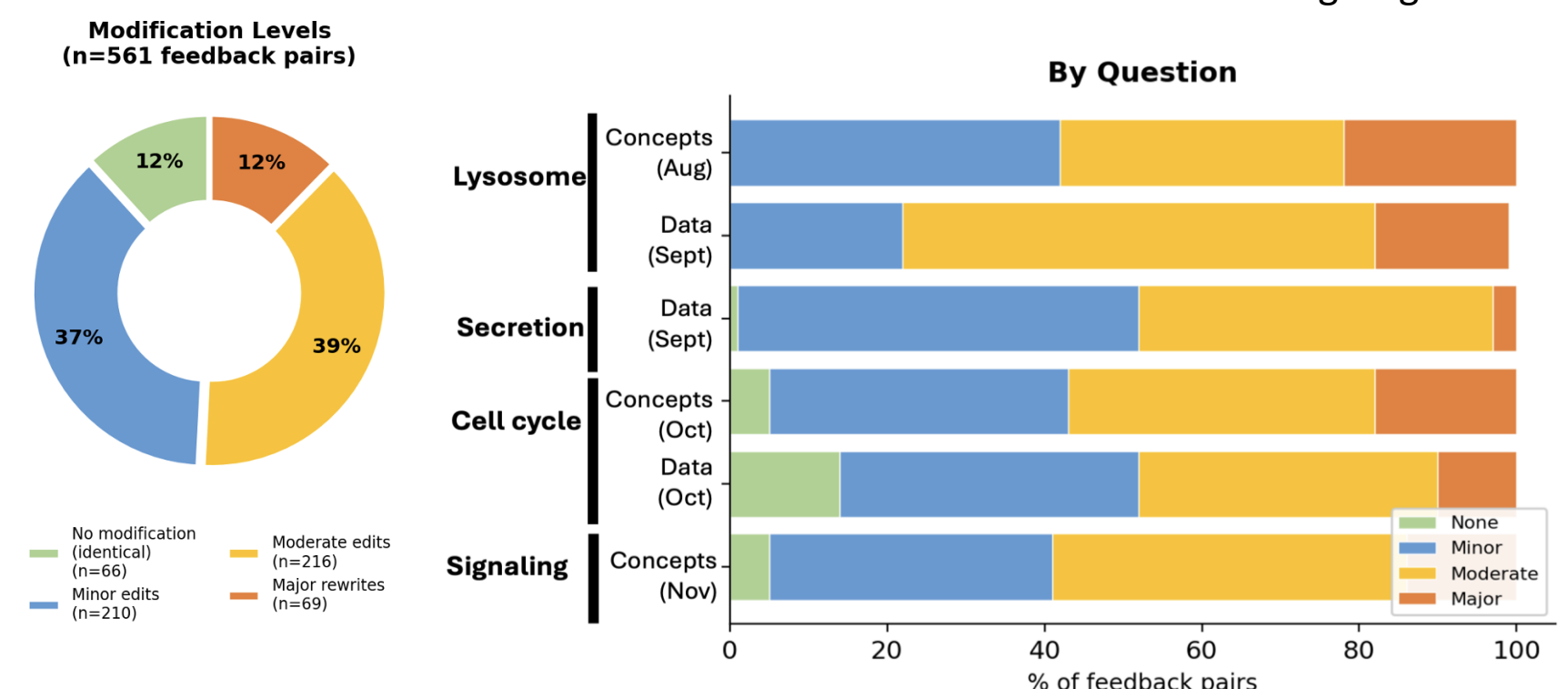


Fig. 5. Extent of revisions to AI-generated feedback across SAQs

The donut chart (left) summarises overall modification levels applied to feedback, categorised as none/identical, minor edits, moderate edits, and major rewrites (counts shown in the legend). The stacked horizontal bars (right) break down the same four modification levels by topics and SAQs in term of percentage of feedback pairs in each category. Colours are consistent across panels (green = none, blue = minor, yellow = moderate, orange = major).

- Most instructor edits targeted **pedagogy, rubric alignment, and clarity of expression** (e.g., restructuring and tone adjustment)
- Factual corrections were uncommon (11%) (Fig. 6).
- Enabled scalable, personalized feedback while preserving instructional judgement through expert review.

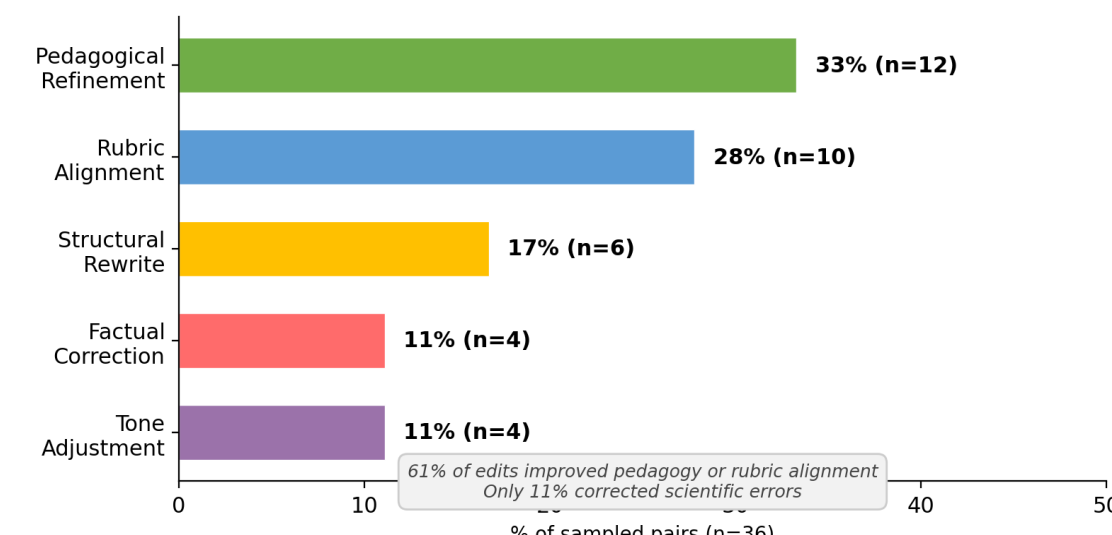


Fig 6. Type of Changes instructors made for AI-generated feedback

Qualitative coding of 36 stratified instructor revision pairs identified 61% of coded edits focused on pedagogy or rubric alignment, while only 11% addressed scientific errors

- Students generally found personalized feedback **useful for clarifying course expectations**
- Students **preferred individualized feedback**
- Minimal concern that the feedback was AI-assisted.**
- Practice SAQs were seen as appropriately challenging, with **strong demand for more** (Fig. 7).

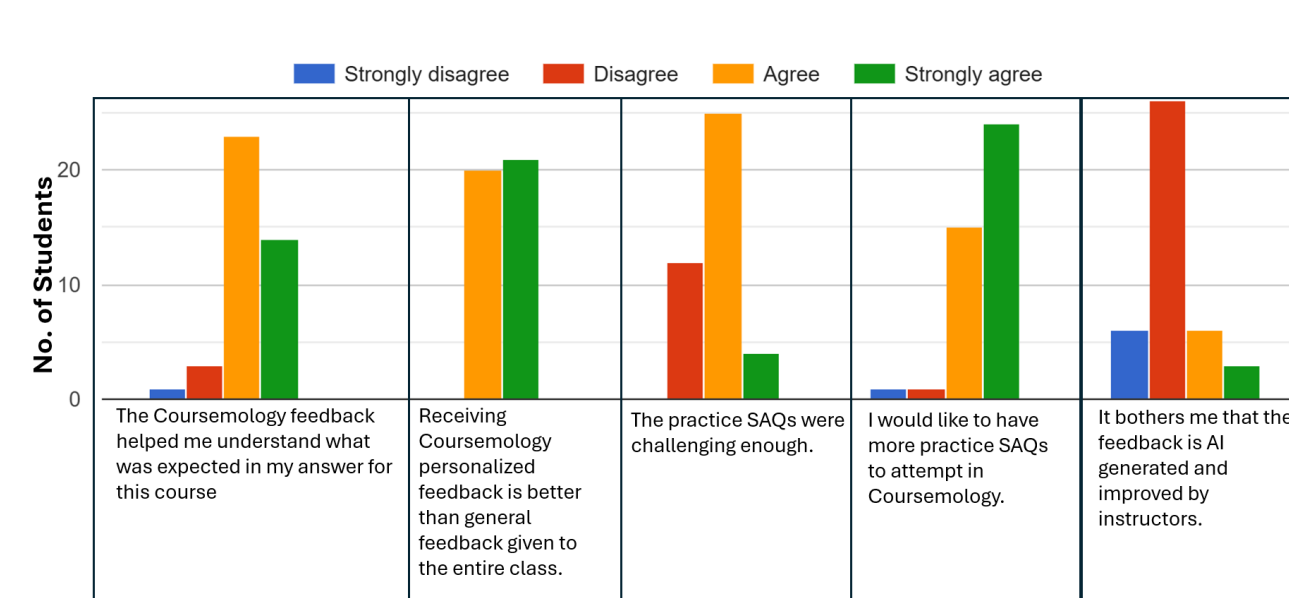


Fig 7. Student perceptions of AI-assisted feedback and practice SAQs.

Distribution of student responses (n = 41) to five survey statements on the use of Coursemology-based AI-generated, instructor-reviewed feedback and practice short-answer questions (SAQs). Responses are shown on a 4-point Likert scale (Strongly disagree, Disagree, Agree, Strongly agree).

## DISCUSSION

- This Gen-AI-assisted, Human-in-the-Loop approach with well-designed prompts, model answers, and instructor review for personalized feedback to student can be implemented across disciplines or institutions.
- Effective integration with a compatible learning platform, prompt design (iteratively refined), data privacy, and staff training are critical for success.
- Other features such as structured rubrics and automated grading functions developed in this work will be evaluated for reliability and alignment in the future.

## REFERENCES

Chua, C. E., Ng, I. K. S., Yuen, K., Teo, D. B., & Luke, N. (2026). Re-imagining discharge summary training through artificial intelligence. *Medical Teacher*, 1–4.

Memarian, B., & Doleck, T. (2024). *Human-in-the-loop in artificial intelligence in education: A review and entity-relationship analysis* *Comput. Hum. Behav. Artif. Hum.* 2(1):100053